



## UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Adress: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
[www.uspto.gov](http://www.uspto.gov)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/610,690	06/30/2003	Charles J. Levine	MSFT-1797 (303687.01)	2925
41505	7590	08/24/2010	EXAMINER	
WOODCOCK WASHBURN LLP (MICROSOFT CORPORATION) CIRA CENTRE, 12TH FLOOR 2929 ARCH STREET PHILADELPHIA, PA 19104-2891			STATE, BRENT S	
ART UNIT	PAPER NUMBER			
		2161		
MAIL DATE	DELIVERY MODE			
08/24/2010	PAPER			

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b> 10/610,690	<b>Applicant(s)</b> LEVINE ET AL.
	<b>Examiner</b> BRENT STACE	<b>Art Unit</b> 2161

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --  
**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
  - If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
  - Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED. (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) Responsive to communication(s) filed on 01 June 2010.  
 2a) This action is FINAL.      2b) This action is non-final.  
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) Claim(s) 1-3,5-15,20 and 21 is/are pending in the application.  
 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.  
 5) Claim(s) \_\_\_\_\_ is/are allowed.  
 6) Claim(s) 1-3,5-15,20 and 21 is/are rejected.  
 7) Claim(s) \_\_\_\_\_ is/are objected to.  
 8) Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) The specification is objected to by the Examiner.  
 10) The drawing(s) filed on 26 October 2007 is/are: a) accepted or b) objected to by the Examiner.  
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).  
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).  
 a) All    b) Some \* c) None of:  
 1. Certified copies of the priority documents have been received.  
 2. Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.  
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- 1) Notice of References Cited (PTO-892)  
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)  
 3) Information Disclosure Statement(s) (PTO/SB/08)  
 Paper No(s)/Mail Date \_\_\_\_\_
- 4) Interview Summary (PTO-413)  
 Paper No(s)/Mail Date \_\_\_\_\_  
 5) Notice of Informal Patent Application  
 6) Other: \_\_\_\_\_

**DETAILED ACTION**

***Remarks***

1. This communication is responsive to the amendment filed June 1<sup>st</sup>, 2010. Claims 1-3, 5-15, and 20-21 are pending. In the amendment filed June 1<sup>st</sup>, 2010, Claims 1, 11, 13, 20, and 21 are amended, Claims 4, 16-19, and 22 are canceled, and Claims 1, 11, 20, and 21 are independent Claims.

***Continued Examination Under 37 CFR 1.114***

2. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 6/1/10 has been entered.

***Response to Arguments***

3. Applicant's arguments filed June 1<sup>st</sup>, 2010 with respect to Claims 11-15 and 20 have been considered but are not persuasive.
4. As to Claims 1-3, 5-10, and 21, amendments to the claims may have allowable subject matter, however, the claims are still rejected under 35 U.S.C. 112 1<sup>st</sup> and/or 2<sup>nd</sup> paragraph. See below.

5. As to the applicant's arguments with respect to exemplarily Claim 11 (including Claim 20) for the prior art(s) allegedly not teaching "**generating different synthetic data from the same data by use of different parameters and seeds**," the examiner respectfully disagrees. Duckworth, p. 2 teaches:

Each time a program is run, the pseudo-random values returned by Rnd are always the same! Obviously, this is very predictable and not random at all. However, VB has also provided the Randomize method.

Use Randomize to initialize VB's random number generator. This is called seeding. To avoid generating the same sequence of pseudo-random values, call Randomize before you call Rnd; but this only has to be done once...

If randomize is not called before Rnd, the same series of pseudo-random values will be returned. It's as if VB always starts returning values from the first entry of that conceptual internal list of 1 million pseudo-random values.

There are occasions when it will be desirable to repeat a series of pseudo-random values, but not always from the "first" entry of the internal list of 1 million. To start somewhere else in that conceptual list, call RND with a negative parameter, followed by a call to Randomize, passing it the starting point. For example:

'Tell VB to initialize using Randomize's parameter  
Rnd -1  
'Tell VB to use 123 as the initialization point (seed)  
Randomize 123

By replacing 123 with a different seed value, the starting point into the conceptual list of 1 million pseudo-random values changes, but each time the program is run, the values returned will always be in the same sequence.

By the above, Duckworth teaches that different synthetic data (output) is generated from the same data (conceptual list of 1 million values) by using different seeds. It should be noted that the seed is a parameter to the Randomize function.

Thus, different synthetic data (output) is generated from the same data (conceptual list of 1 million values) by using different parameters and seeds.

6. As to the applicant's arguments with respect to exemplarily Claim 11 (including Claim 20) for the prior art(s) allegedly not teaching "**data sets that each differ by use of different seeds and parameters for generating the data sets**," the examiner respectfully disagrees. Generally, this argument appears to have been met above. Note that the Rnd function/code can be utilized many times to produce random numbers in a VB program.

7. Any other claims argued merely because of a dependency on a previously argued claim(s) in the arguments presented to the examiner, dated June 1<sup>st</sup>, 2010, are moot in view of the examiner's interpretation of the claims and art and are still considered rejected based on their respective rejections from at least a prior Office action (part(s) of recited again below).

***Response to Amendment***

***Claim Objections***

8. Claims 1 and 21 are objected to because of the following informalities:

a. Further clarity is suggested for Claims 1 and 21 regarding what is input into the "deterministic data generation module." For instance, Claim 1, line 8 could recite "accepting, as a first input to a deterministic data generation module, at least one of." Additionally, the generated first and second synthetic data does not appear to be an output, which is the general concept of the invention. As a

whole it is unclear what is input and output to and from the deterministic data generation module. The applicants are invited to schedule a telephone interview with the examiner regarding clarifying amendments to these claims.

Appropriate correction is required.

***Claim Rejections - 35 USC § 112***

9. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

10. Claims 1-3, 5-10, and 21 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The examiner could not find sufficient support in the specification for the amendments to the claims.

11. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

12. Claims 9 and 10 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Claims 9 and 10 both recite "the synthetic data" when Claim 1 (from which both Claims 9 and 10 depend from) recite a first synthetic

data and a second synthetic data. As such, it is unclear what synthetic data (first, second or both first and second) the claim is referring to.

***Claim Rejections - 35 USC § 102***

13. In light of the applicant's respective arguments or respective amendments, the previous 35 USC § 102 rejections to the claims have been withdrawn.

***Claim Rejections - 35 USC § 103***

14. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

15. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

16. Claims 11-15 and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over "Random Numbers" (Duckworth) in view of "Quickly Generating Billion-Record Synthetic Databases" (Gray et al.).

For **Claim 11**, Duckworth teaches: "A computer-implemented method for generating data [Duckworth, p. 1, line under "How Do I Get Random Numbers?" generating random numbers] comprising:

- providing by at least one computer processor a deterministic data generation module stored on at least one medium, [Duckworth, p. 1, line under "How Do I Get Random Numbers?" Rnd is a data generation module and generating random numbers in VB requires media holding the instructions] the deterministic data generation module accepting inputs for processing [Duckworth, p. 2, "Repeating Random Numbers", one input required (or defaulted) is the seed another is the long list] to generate a plurality of data sets, [Duckworth, p. 2, "Repeating Random Numbers", each Rnd called without a new seed ("values returned will...be...the same sequence")] each data set having synthesized data [Duckworth, p. 1, "Why Psuedo?" (the indented subsection), random numbers are synthesized data] wherein within the data set each data element has a sequence number, [Duckworth, p. 1, "Why Psuedo?" (indented subsection), starting point (e.g. first value and 105,768<sup>th</sup> value) acting as indexes] and each data set is organized such that the data is positioned from lowest sequence number to highest sequence number in a sequential fashion, [Duckworth, p. 1, "Why Psuedo?" (indented subsection), a long (1 million) list has fixed

positions/sequence numbers that, how you view it will be in order from lowest to highest] and wherein the synthesized data of each data set differs by using different seeds and parameters to generate the data sets; [Duckworth, p. 1, "Why Psuedo?" (indented subsection), long list doesn't change, only starting point/seed does. It should be noted that the seed is a parameter to the Randomize function]

- providing by at least one computer processor each of the seeds as input to the deterministic data generation module, [Duckworth, p. 2, "Repeating Random Numbers", one input into Rnd required (or defaulted) is the seed. Seed can be set using Randomize] each of the seeds acting to position the deterministic data generation module to regenerate data having a predefined sequence number, [Duckworth, p. 2, "Repeating Random Numbers", reading the list @ the same seed place will regenerate same data] wherein the seed value is derived from the predefined sequence number, [Duckworth, p. 2, "Repeating Random Numbers" with Duckworth, p. 1, "Why Psuedo?" (the indented subsection), predefined sequence number is the seed] and wherein the sequence number represents a starting point [Duckworth, p. 2, "Repeating Random Numbers", seed in starting position in long list]...wherein each of the seeds is within a range allowed by at least one parameter of the data generation module, [Duckworth, p. 2, "Repeating Random Numbers", list of one million (range is 1 to 1million), 123 is within the range] wherein the at least one parameter is configured to cause the data generation module to generate synthetic data which is repeatable; [Duckworth, p.

2, "Repeating Random Numbers", reading the list @ the same seed place will regenerate same data] and."

Duckworth discloses the above limitations but does not explicitly teach:

- "...from which the synthetic data is used as input to process whose performance is to be evaluated
- ...schematizing the synthesized data according to a predefined data schema definition."

With respect to Claim 11, an analogous art, Gray, teaches:

- "...from which the synthetic data is used as input to process whose performance is to be evaluated [Gray, p. 243, Abstract, evaluating database systems, databases have at least a process for, for example, retrieving data]
- ...schematizing the synthesized data according to a predefined data schema definition" [Gray, p. 244, paragraph under section "3. Sequential database generation," when data generated it is put into the database, a data schema of the database must exists in order to know where to store the data (e.g. tables)].

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Gray and Duckworth before him/her to combine Gray with Duckworth because both inventions are directed towards generating random data.

Gray's invention would have been expected to successfully work well with Duckworth's invention because both inventions use random data generation. Duckworth discloses about the Rnd and Randomize function in VB comprising generating random data. However, Duckworth does not expressly disclose that data

can be randomly generated in a parallel fashion or generating random data for a database having a schema for use in benchmarking nor schematizing the synthesized data according to a predefined data schema definition. Gray discloses quickly generating billion-record synthetic databases (title) comprising using parallel functions to generate the data and inputting the data in to a database having a schema to evaluate database system performance.

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Gray and Duckworth before him/her to take the parallel functions and intended use from Gray and install it into the teachings of Duckworth, thereby offering the obvious advantage of parallelism to get generation speedup and scaleup (Gray, p. 243, Abstract) and being able to use the data for the intended use of evaluating database system performance.

**Claim 12** can be mapped to Duckworth (as modified by Gray) as follows: "The computer-implemented method as recited in claim 11, further comprising communicating the synthesized data to cooperating data environments" [Duckworth, p. 2, "Repeating Random Numbers", random numbers are communicated to the computer and program].

**Claim 13** can be mapped to Duckworth (as modified by Gray) as follows: "The computer-implemented method as recited in claim 11, further comprising changing the value of each of the seeds" [Duckworth, p. 2, "Repeating Random Numbers", seed is changed using Randomize].

**Claim 14** can be mapped to Duckworth (as modified by Gray) as follows: "The computer-implemented method as recited in claim 11, processing the synthesized data by cooperating environments as part of a benchmarking study" [Gray, p. 243, Abstract, evaluating database systems].

**Claim 15** can be mapped to Duckworth (as modified by Gray) as follows: "The computer-implemented method as recited in claim 11, further comprising schematizing the synthesized data according to a predefined data schema definition" [Gray, p. 243, Title, Abstract, and Introduction, Gray is concerned with "generating billion-record synthetic databases" databases have predefined data schema definition].

For **Claim 20**, Duckworth teaches: "A method to generate repeatable synthesized data [Duckworth, p. 1, "Why Psuedo?" with Duckworth, p. 2, "Repeating Random Numbers", random synthesized numbers can be repeatable on same seed/startng position] comprising:

- executing by at least one computer processor a deterministic data generation function to generate a plurality of data sets [Duckworth, p. 2, "Repeating Random Numbers", each Rnd called without a new seed ("values returned will...be...the same sequence")] corresponding to sequential numbers, [Duckworth, p. 1, "Why Psuedo?" (indented subsection), starting point (e.g. first value and 105,768<sup>th</sup> value) acting as indexes] the numbers associated with a data element of each data set, [Duckworth, p. 1, "Why Psuedo?" (indented subsection), starting point (e.g. first value and 105,768<sup>th</sup> value) acting as indexes] wherein each data element and associated number are different in each data set wherein the data

- sets are generated by different seeds and parameters; [Duckworth, p. 1, "Why Psuedo?" (indented subsection), long list doesn't change, only starting point/seed does. It should be noted that the seed is a parameter to the Randomize function]
- setting by at least one computer processor one of the seeds of a corresponding data set to act as input for the deterministic data generation function [Duckworth, p. 2, "Repeating Random Numbers", Rnd is a data generation function, one input required (or defaulted) is the seed. Seed can be set using Randomize] such that the input drives the deterministic data generation function to regenerate data corresponding to a particular sequential number, [Duckworth, p. 2, "Repeating Random Numbers", reading the list @ the same seed place will regenerate same data] wherein one of the seeds seed is within a range allowed by at least one parameter of the data generation function, [Duckworth, p. 2, "Repeating Random Numbers", list of one million (range is 1 to 1million), 123 is within the range] wherein the at least one parameter is configured to cause the data generation function to generate synthetic data which is repeatable and different for each data set, [Duckworth, p. 2, "Repeating Random Numbers", reading the list @ the same seed place will regenerate same data] wherein the seed is set for each discrete data element that may be regenerated; [Duckworth, p. 2, "Repeating Random Numbers", seed is set @ the beginning for each random number requested. Alternatively, the seed can be changed using Randomize] and."
- Duckworth discloses the above limitations but does not explicitly teach:

- "...testing performance of a system by providing said data set as input to said system and measuring behavior of said system using said data set."

With respect to Claim 20, an analogous art, Gray, teaches:

- "...testing performance of a system by providing said data set as input to said system and measuring behavior of said system using said data set" [Gray, p. 243, Abstract, evaluating database systems].

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Gray and Duckworth before him/her to combine Gray with Duckworth because both inventions are directed towards generating random data.

Gray's invention would have been expected to successfully work well with Duckworth's invention because both inventions use random data generation. Duckworth discloses about the Rnd and Randomize function in VB comprising generating random data. However, Duckworth does not expressly disclose that data can be randomly generated in a parallel fashion or generating random data for a database having a schema for use in benchmarking. Gray discloses quickly generating billion-record synthetic databases (title) comprising using parallel functions to generate the data and inputting the data in to a database having a schema to evaluate database system performance.

It would have been obvious to one of ordinary skill in the art at the time of invention having the teachings of Gray and Duckworth before him/her to take the parallel functions and intended use from Gray and install it into the teachings of Duckworth, thereby offering the obvious advantage of parallelism to get generation

Art Unit: 2161

speedup and scaleup (Gray, p. 243, Abstract) and being able to use the data for the intended use of evaluating database system performance.

***Conclusion***

17. Any prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Applicant is advised that, although not used in the rejections above, prior art cited on any PTO-892 form and not relied upon is considered materially relevant to the applicant's claimed invention and/or portions of the claimed invention.
18. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brent S. Stace whose telephone number is 571-272-8372 and fax number is 571-273-8372. The examiner can normally be reached on M-F 10am-6:30pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Apu M. Mofiz can be reached on 571-272-4080. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

/BRENT STACE/  
Examiner, Art Unit 2161